Money Laundering: Using a hybrid approach to detect illicit shell companies

Tiwari, Milind; Gepp, Adrian; Kumar, Kuldeep

ABSTRACT

Purpose: This paper aims to detect illicit shell companies by using a combination of graph algorithms and modern supervised learning that are trained on prior cases of money laundering.

Methodology: Data on prior cases of illicit companies involved in money laundering, as provided by Transparency International, are collected through OpenCorporates, cleansed through OpenRefine and transported to the graph database platform Neo4J. Several graph algorithms were used to obtain measures of community, similarity and centrality. The resulting scores were then used by decision-tree based models to classify the companies that are involved in illicit behaviour.

Findings: By using a hybrid approach comprising graphical network analysis and supervised learning with trees, our modelling correctly classified illicit companies 96.10% of the time and fraudulent companies 90.36% of the time on the training data. Importantly, these strong results were maintained on new test data. The classification accuracy for illicit and licit companies was 94.81% and 90.96%, respectively, on the test data.

Research Implications: The accuracy of the models documented in the study provides support for using a combination of graph analytics and modern supervised learning in aiding detection efforts.

Practical Implications: This paper takes steps in strengthening the first line of defence, that is, incorporation services and corporate registries, in the fight against money laundering by proposing a network topology for identifying hidden links among companies and using a hybrid approach to detect illicit entities.

Originality: No studies in the past have considered companies and their relationships from multiple cases in the same graph network for using a hybrid approach to detect illicit companies.

Keywords: Graph Database; Money Laundering; Neo4J; Network Analysis; Shell Companies; Supervised Learning